**Final Part 2**

This assignment is OPTIONAL and can only help your grade!

There are 4 questions. For each question, you can earn UP TO 2 points on your Part 1 Final grade. *Example: You got a 72% on Part 1. You could raise your score to a 74% if you do 1 question. You could raise your score to an 80% if you do 4 questions. (You cannot get higher than a 100 though)*

The data for these questions are in FinalPart2.csv. Here is a list explaining the variables. All variables are scale/interval.

IncomeC = Income during college

HealthC = Health rating during college

StressC = Stress during college

LifeSatC = Life satisfaction during college

Income10 = Income 10 years after graduation

Health10 = Health 10 years after graduation

Stress10 = Stress 10 years after graduation

LifeSat10 = Life 10 years after graduation

**Question 1**

A. Run a correlation matrix with all 4 college variables. Copy and paste the correlation matrix below. Write a paragraph that summarizes the results of the correlation table. Report each correlation in APA format. For each correlation, describe the strength, direction, and statistical significance and write a sentence summarizing the relationship between the variables (e.g. “as \_\_\_ goes up, so does \_\_\_\_”).

**Correlation Matrix**

| **Pearson Correlations** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | | **IncomeC** | | **StressC** | | **HealthC** | | **LifeSatC** | |
| IncomeC |  | Pearson's r |  | — |  |  |  |  |  |  |  |
| p-value |  | — |  |  |  |  |  |  |  |
| StressC |  | Pearson's r |  | -0.505 |  | — |  |  |  |  |  |
| p-value |  | 0.023 |  | — |  |  |  |  |  |
| HealthC |  | Pearson's r |  | 0.583 |  | -0.725 |  | — |  |  |  |
| p-value |  | 0.007 |  | < .001 |  | — |  |  |  |
| LifeSatC |  | Pearson's r |  | 0.602 |  | -0.308 |  | 0.273 |  | — |  |
| p-value |  | 0.005 |  | 0.187 |  | 0.245 |  | — |  |
|  | | | | | | | | | | | |
|  | | | | | | | | | | | |

In the correlation matrix with the 4 college variables, stress and income were found to have a moderately negative correlation, *r*=-0.505, *p*= 0.023. This meaning that as income went up stress went down. Income and health were found to have a moderately positive correlation, *r*=0.583, *p*=0.007. This means that as income went up health went up. Life satisfaction and income were found to have a strong positive correlation *r=*0.602, *p=*0.005, This means that as income went up life satisfaction also went up. It was found that health and stress were strongly negatively correlated, *r=-*0.725, *p*=<.001. This means that as health went up stress went down. It was found that health and life satisfaction had a weak positive correlation, *r*=0.273, *p*=0.245. This means that as health went up so did life satisfaction. Finally stress and life satisfaction were found to have a moderately weak negative correlation, *r*=-0.308, *p*=0.187. This means that as stress goes up life satisfaction goes down.

B. Run a correlation matrix with all 4 post-graduation variables. Copy and paste the correlation matrix below. For each *significant correlation,* describe the strength, direction, and statistical significance in APA format and write a sentence summarizing the relationship between the variables (e.g. “as \_\_\_ goes up, so does \_\_\_\_”).

**Correlation Matrix**

| **Pearson Correlations** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | | **Income10** | | **Health10** | | **Stress10** | | **LifeSat10** | |
| Income10 |  | Pearson's r |  | — |  |  |  |  |  |  |  |
| p-value |  | — |  |  |  |  |  |  |  |
| Health10 |  | Pearson's r |  | 0.256 |  | — |  |  |  |  |  |
| p-value |  | 0.276 |  | — |  |  |  |  |  |
| Stress10 |  | Pearson's r |  | -0.098 |  | -0.478 |  | — |  |  |  |
| p-value |  | 0.682 |  | 0.033 |  | — |  |  |  |
| LifeSat10 |  | Pearson's r |  | 0.551 |  | 0.121 |  | -0.124 |  | — |  |
| p-value |  | 0.012 |  | 0.611 |  | 0.603 |  | — |  |
|  | | | | | | | | | | | |

It was found that income and health had a weak positive correlation, *r*=0.256, *p*=0.276. This means that as income went up so did health. There was a very weak negative correlation between income and stress, *r=*-0.098, *p*=0.682. As income got higher stress went down. It was found that health and stress are moderately negatively correlated, *r=*-0.478, *p*=0.033. As stress goes up, health goes down. It was found that income and life satisfaction are moderately/strongly positively correlated, *r=*0.551, *p*=0.012. As income goes up, life satisfaction also goes up. Health and life satisfaction have a weak positive correlation, *r*=0.121, *p*=0.603. As health goes up so does life satisfaction. Finally stress and life satisfaction have a weak negative correlation, *r=*-0.124, *p*=0.603. This means that as stress goes up, life satisfaction goes down.

**Question 2**

Run a multiple regression with the 4 college variables predicting Income at 10 years after college. Paste the tables below. Write a paragraph explaining the results. Report the R2 and *F* ratio for the full multiple regression. Include what the R2 means. For each predictor, report *t* test results/ statistical significance using APA format. Make sure to summarize the findings (what is the take home message?). Lastly, write the full regression equation (You can use the symbols or variable names).

**Linear Regression**

| **Model Summary** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | | **R** | | **R²** | | **Adjusted R²** | | **RMSE** | | | |
| 1 |  | 0.682 |  | 0.465 |  | 0.322 |  | 21.488 | | |  |
|  | | | | | | | | | | | |
| **ANOVA** | | | | | | | | | | | | | | | | |
| **Model** | |  | | | **Sum of Squares** | | | | **df** | | **Mean Square** | | **F** | | **p** | |
| 1 |  | Regression | |  | 6010.428 | | |  | 4 |  | 1502.607 |  | 3.254 |  | 0.041 |  |
|  |  | Residual | |  | 6925.772 | | |  | 15 |  | 461.718 |  |  |  |  |  |
|  |  | Total | |  | 12936.200 | | |  | 19 |  |  |  |  |  |  |  |
|  | | | | | | | | | | | | | | | | |

| **Coefficients** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | |  | | **Unstandardized** | | **Standard Error** | | **Standardized** | | **t** | | **p** | |
| 1 |  | (Intercept) |  | -57.303 |  | 66.509 |  |  |  | -0.862 |  | 0.402 |  |
|  |  | IncomeC |  | 1.702 |  | 0.667 |  | 0.722 |  | 2.551 |  | 0.022 |  |
|  |  | HealthC |  | 1.283 |  | 1.170 |  | 0.327 |  | 1.096 |  | 0.290 |  |
|  |  | StressC |  | 5.302 |  | 2.339 |  | 0.632 |  | 2.267 |  | 0.039 |  |
|  |  | LifeSatC |  | -0.342 |  | 0.540 |  | -0.152 |  | -0.634 |  | 0.536 |  |
|  | | | | | | | | | | | | | |

A multiple linear regression was calculated to predict the income 10 years after college based on income during college, health during college, stress during college, and life satisfaction in college. The regression equation was significant (*F*(4,15)= 3.254, *p*= 0.041), with an R2 of 0.465. These variables account for 46.5% of the variability in income 10 years after college. The analysis shows that income in college is a significant predictor of income 10 years after graduation (*b=* 1.702, *t*(19)=2.551, *p=*0.022). This shows that the income 10 years after graduation will increase by 2.551. The analysis shows that health during college was not a strong predictor of income 10 years after graduation (*b=* 1.283, *t*(19)=1.096, *p=*0.290) and neither was life satisfaction in college (*b=* -0.342, *t*(19)=-0.634, *p=*0.536). However stress in college did prove to be a significant predictor of income 10 years after graduation (*b=* 5.302, *t*(19)=2.267, *p=*0.039). For the increase in stress rating income increased by 5.302.

ŷ= (1.702)(IncomeC) + (1.283)(HealthC) + (5.302)(StressC) + (-0.342)(LifeSatC)-57.303

**Question 3**

Run a multiple regression with the 4 college variables predicting stress at 10 years after college. Paste the tables below. Write a paragraph explaining the results. Report the R2 and *F* ratio for the full multiple regression. Include what the R2 means. For each predictor, report *t* test results/ statistical significance using APA format. Make sure to summarize the findings (what is the take home message?). Lastly, write the full regression equation (You can use the symbols or variable names).

**Linear Regression**

| **Model Summary** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | | **R** | | **R²** | | **Adjusted R²** | | **RMSE** | | | |
| 1 |  | 0.753 |  | 0.566 |  | 0.451 |  | 1.545 | | |  |
|  | | | | | | | | | | | |
| **ANOVA** | | | | | | | | | | | | | | | | |
| **Model** | |  | | | **Sum of Squares** | | | | **df** | | **Mean Square** | | **F** | | **p** | |
| 1 |  | Regression | |  | 46.762 | | |  | 4 |  | 11.691 |  | 4.900 |  | 0.010 |  | |
|  |  | Residual | |  | 35.788 | | |  | 15 |  | 2.386 |  |  |  |  |  | |
|  |  | Total | |  | 82.550 | | |  | 19 |  |  |  |  |  |  |  | |
|  | | | | | | | | | | | | | | | | |

| **Coefficients** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | |  | | **Unstandardized** | | **Standard Error** | | **Standardized** | | **t** | | **p** | |
| 1 |  | (Intercept) |  | 3.454 |  | 4.781 |  |  |  | 0.722 |  | 0.481 |  |
|  |  | IncomeC |  | 0.039 |  | 0.048 |  | 0.209 |  | 0.822 |  | 0.424 |  |
|  |  | HealthC |  | -0.045 |  | 0.084 |  | -0.145 |  | -0.539 |  | 0.598 |  |
|  |  | StressC |  | 0.499 |  | 0.168 |  | 0.744 |  | 2.967 |  | 0.010 |  |
|  |  | LifeSatC |  | 0.013 |  | 0.039 |  | 0.074 |  | 0.344 |  | 0.736 |  |
|  | | | | | | | | | | | | | |

A multiple linear regression was calculated to predict stress 10 years after college graduation, based on income during college, health during college, stress during college, and life satisfaction during college. The regression was significant (*F*(4,15)= 4.900, *p*= 0.010), with an R2 of 0.566. These variables account for 56.6% of variability in stress 10 years after graduation. The analysis shows that income during college was not a predictor of stress 10 years after graduation (*b=* 0/039, *t*(19)=0.822, *p=*0.424). Health during college also was not a predictor of stress 10 years after graduation (*b=* -0.045, *t*(19)=-0.539, *p=*0.598). Life satisfaction in college also was not a predictor for stress 10 years after graduation (*b=* 0.013, *t*(19)=0.344, *p=*0.736).Stress in college was a predictor for stress 10 years after graduation (*b=* 0.499, *t*(19)=2.967, *p=*0.010) this meant that the stress levels 10 years after graduation would be representative of their stress levels during college.

ŷ= (0.039)(IncomeC) + (-0.045)(HealthC) + (0.499)(StressC) + (0.013)(LifeSatC)-3.454

**Question 4**

Run a multiple regression with stress during college and income during college predicting health at 10 years after college. Paste the tables below. Write a paragraph explaining the results. Report the R2 and *F* ratio for the full multiple regression. Include what the R2 means. For each predictor, report *t* test results/ statistical significance using APA format. Make sure to summarize the findings (what is the take home message?). Lastly, write the full regression equation (You can use the symbols or variable names).

## Linear Regression

| **Model Summary** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | | **R** | | **R²** | | **Adjusted R²** | | **RMSE** | | | |
| 1 |  | 0.735 |  | 0.541 |  | 0.418 |  | 5.995 | | |  |
|  | | | | | | | | | | | |
| **ANOVA** | | | | | | | | | | | | | | | | |
| **Model** | |  | | | **Sum of Squares** | | | | **df** | | **Mean Square** | | **F** | | **p** | |
| 1 |  | Regression | |  | 634.715 | | |  | 4 |  | 158.679 |  | 4.415 |  | 0.015 |  |
|  |  | Residual | |  | 539.085 | | |  | 15 |  | 35.939 |  |  |  |  |  |
|  |  | Total | |  | 1173.800 | | |  | 19 |  |  |  |  |  |  |  |
|  | | | | | | | | | | | | | | | | |

| **Coefficients** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | |  | | **Unstandardized** | | **Standard Error** | | **Standardized** | | **t** | | **p** | |
| 1 |  | (Intercept) |  | 25.565 |  | 18.556 |  |  |  | 1.378 |  | 0.188 |  |
|  |  | IncomeC |  | 0.166 |  | 0.186 |  | 0.233 |  | 0.890 |  | 0.387 |  |
|  |  | HealthC |  | 0.458 |  | 0.326 |  | 0.388 |  | 1.403 |  | 0.181 |  |
|  |  | StressC |  | -0.741 |  | 0.653 |  | -0.293 |  | -1.135 |  | 0.274 |  |
|  |  | LifeSatC |  | -0.123 |  | 0.151 |  | -0.181 |  | -0.815 |  | 0.428 |  |
|  | | | | | | | | | | | | | |

A multiple linear regression was calculated to predict health 10 years after college graduation, based on income during college, health during college, stress during college, and life satisfaction during college. The regression was significant (*F*(4,15)= 4.415, *p*= 0.015), with an R2 of 0.541. These variables account for 54.1% of variability in health 10 years after graduation. None of the variables proved to be good predictors of health 10 years after graduation. Income in college was not a significant predictor (*b=* 0.166, *t*(19)=-0.890, *p=*0.387). Health during college was not an accurate predictor of health 10 years after graduation (*b=* 0.458, *t*(19)=1.403, *p=*0.181), neither was stress in college (*b=* -0.741, *t*(19)=-1.135, *p=*0.274), or life satisfaction in college (*b=* -0.123, *t*(19)=-0.815, *p=*0.428). This shows that there are better indicators of health 10 years after graduation than income during college, health during college, stress during college, and life satisfaction during college.

ŷ= (0.166)(IncomeC) + (0.458)(HealthC) + (-0.741)(StressC) + (-0.123)(LifeSatC)-25.565